

Mikrozensus Scientific Use File 2001: Dokumentation und Datenaufbereitung

Shahla, Hossein

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Shahla, H. (2004). *Mikrozensus Scientific Use File 2001: Dokumentation und Datenaufbereitung*. (ZUMA-Methodenbericht, 2004/08). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-48506-7>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

ZUMA-Methodenbericht 2004/08

Mikrozensus Scientific Use File 2001
Dokumentation und Datenaufbereitung

Hossein Shahla

Oktober 2004

ISSN 1610-9953

ZUMA
Quadrat B2,1
Postfach 12 21 55
68072 Mannheim
Telefon: (0621) 1246 - 253
Telefax: (0621) 1246 - 100
E-mail: shahla@zuma-mannheim.de

Inhaltsverzeichnis

1.	Einleitung	1
2.	Mikrozensus 2001	1
2.1	Kurzbeschreibung.....	1
2.2	Datenerhebung	2
2.3	Die Stichprobe.....	2
2.4	Das Erhebungsprogramm.....	3
3.	Der faktisch anonymisierte Mikrozensus – Das Scientific Use File –.....	4
4.	Aufbereitung des Rohdatenfiles zum Mikrozensus Scientific Use File 2001 und Überprüfung der Plausibilität des erstellten Systemfiles	5
4.1	Einlesen des Rohdatenmaterials und Erstellung des Systemfiles	5
4.2	Gewichtung und Hochrechnung zum Vergleich mit den Veröffentlichungen des Statistischen Bundesamtes	6
4.3	Bevölkerungsbegriffe und Abgrenzungen im Mikrozensus.....	7
4.4	Plausibilitätsprüfung des aufbereiteten Mikrozensus Scientific Use Files 2001	8
Anhang:		
A)	Randverteilung ausgewählter Variablen des Scientific Use Files 2001.....	10
B)	SPSS-Syntax für die Plausibilitätsprüfungen des Scientific Use Files 2001	14
C)	Weiterführende Literaturhinweise zum Mikrozensus	16

1. Einleitung

Mit dem vorliegenden ZUMA-Methodenbericht setzt das German Microdata Lab (GML) die Reihe der Dokumentationen der faktisch anonymisierten Mikrozensusdaten (Scientific Use Files) fort.

Die Bereitstellung der faktisch anonymisierten Daten als Rohdatenfile in ASCII-Format erfolgt durch das Statistische Bundesamt. In Kooperation mit dem Statistischen Bundesamt stellt das German Microdata Lab dem Nutzer ausführliche Informationen über die Scientific Use Files zur Verfügung und bietet Setups zur Aufarbeitung des Datenmaterials und zur Erstellung von System-Files in Datenformaten gängiger Softwarepakete an. Ausführlichere Informationen über die Mikrozensus Scientific Use Files (vorhandene Variablen, verwendete Systematiken, Fragebogen, Schlüsselverzeichnis, intertemporale Vergleichbarkeit, Mikrodaten Tools etc.) werden auch in ZUMA-Methodenberichten zum intertemporalen Vergleich von Mikrozensen¹ und im World Wide Web des German Microdata Lab² bereitgestellt und können dort abgerufen werden.

Im Folgenden wird nach einer kurzen Beschreibung des Mikrozensus und des Mikrozensus Scientific Use Files 2001 der Vorgang der Datenaufbereitung beschrieben und die notwendigen Schritte zur Überprüfung der Plausibilität von eingelesenen Daten erläutert.

2. Mikrozensus 2001

2.1 Kurzbeschreibung

Der Mikrozensus ist mit einem Auswahlsatz von 1% der Bevölkerung die größte jährliche Haushaltsbefragung in Europa. Sie wird seit 1957 in Westdeutschland und seit 1991 in den neuen Bundesländern durchgeführt. Grundlage für die ausführlichen Angaben zu bevölkerungs- und arbeitsmarkt-statistischen Strukturdaten des Mikrozensus ist die gesamte wohnberechtigte Bevölkerung in Deutschland, zu der im April 2001 rund 82 Mio. Menschen gehörten. Der Mikrozensus ist durch eine hohe Kontinuität im Frageprogramm und aufgrund der Auskunftspflicht durch einen niedrigen Unit-Nonresponse gekennzeichnet. Der Mikrozensus ermöglicht differenzierte Analysen auf Personen-, Familien- und Haushalts-

¹ Vgl. ZUMA-Methodenberichte 2004/09, 2003/08 und 2003/02 sowie die ZUMA-Technischen Berichte 2000/08, 2000/09 und 1998/09.

² Vgl.: http://www.gesis.org/Dauerbeobachtung/Mikrodaten/Service_Guide/index.htm

ebene sowohl in Quer- als auch in Längsschnitt und eignet sich für die Beobachtung von Strukturveränderungen sowie zu tiefgegliederten Analysen.

2.2 *Datenerhebung*

Die Angaben im Mikrozensus 2001 beziehen sich auf die Berichtswoche vom 23. bis 29. April 2001. Die Daten des Mikrozensus werden durch die Statistischen Landesämter erhoben. Die Befragungen finden i.d.R. mündlich statt, ein kleiner Teil erfolgt mittels schriftlicher Auskunft der Befragten. Für alle Beteiligten besteht gemäß der gesetzlichen Regelungen Auskunftspflicht. Nur wenige Fragen sind freiwillig zu beantworten. Der Mikrozensus ist eine Haushaltsstichprobe, d.h. alle Personen in einem Haushalt werden entweder direkt befragt, bzw. eine erwachsene Person gibt Auskunft über die weiteren (alle minderjährigen) Haushaltsmitglieder. Der Mikrozensus umfasst auch ausländische Haushalte. Zur Erhebung ihrer Angaben wird der gleiche Fragebogen verwendet wie bei deutschen Haushalten. Die Daten werden von den Statistischen Landesämtern aufbereitet und beim Statistischen Bundesamt zu einem Datensatz zusammengespült.

2.3 *Die Stichprobe*

Für den Mikrozensus 2001 gilt das Stichprobendesign von 1990. Die Stichprobe wurde aus dem Datenmaterial der Volkszählung 1987 gezogen. Das stichprobenmethodische Grundkonzept des Mikrozensus ist die einstufige Klumpenstichprobe. Die Klumpen sind Auswahlbezirke, die ganze Gebäude oder Gebäudeteile umfassen. Zur Verbesserung der Regionalisierbarkeit wurde 1990 die durchschnittliche Größe der Auswahlbezirke verkleinert. Ein Auswahlbezirk umfasst durchschnittlich 9 Wohnungen (vor 1990: 23 Wohnungen). Die Stichprobe wird jährlich durch eine Ergänzungsauswahl der Bautätigkeitsstatistik aktualisiert.

Während für das frühere Bundesgebiet die Grundausswahl auf der Basis der Volkszählung 1987 vorgenommen werden konnte, war für die neuen Bundesländer die Volkszählung 1981 als Auswahlgrundlage zu alt. Aus diesem Grund wurde in den neuen Bundesländern und Ost-Berlin das „Bevölkerungsregister Statistik“ in einem weitgehend analogen Verfahren zur Abgrenzung der Auswahlseinheiten herangezogen. Im Mikrozensus 2001 sind Informationen zu dem Auswahlbezirk, der Gebäudeschicht, der Bogenart und der Zahl der Haushalte enthalten.

2.4 Das Erhebungsprogramm

Das Fragenprogramm des Mikrozensus besteht aus einem gleichbleibenden Grundprogramm, das sozioökonomische Grundinformationen liefert und die laufende Beobachtung des Arbeitsmarktes ermöglicht. Im Grundprogramm, das sich an alle Haushalte richtet, sind im Mikrozensus 2001 folgende Bereiche enthalten:

- Fragen zur Wohnung/zum Haushalt
- Fragen zu den Personen im Haushalt
- Schulbesuch
- Erwerbsbeteiligung
- Erste gegenwärtige Erwerbstätigkeit
- Zweite Erwerbstätigkeit
- Arbeitsuche/Arbeitsplatzwechsel
- Altersvorsorge
- Ausbildung
- Unterhalt, Einkommen

Neben den jährlich zu erhebenden Merkmalen des Grundprogramms (Auswahlsatz 1%) werden in jährlichen Ergänzungsprogrammen mit einem Auswahlsatz von 0,5% weitere Themen erfragt.

Im Ergänzungsprogramm 2001 sind folgende zusätzliche Bereiche enthalten:

- Berufliche und allgemeine Aus- und Fortbildung
- Ergänzende Angaben zur Erwerbstätigkeit
- Frühere Erwerbstätigkeit
- Erwerbsbeteiligung ein Jahr vor der Erhebung

Darüber hinaus gibt es ein Zusatzprogramm mit Merkmalen, die nur im Abstand von 4 Jahren erhoben werden (Auswahlsatz bis zu 0,5%). Im Rahmen der vierjährigen Zusatzprogramme enthält der Mikrozensus 2001 folgende Angaben:

- Zusatzangaben zur Erwerbstätigkeit
- Private und betriebliche Altersvorsorge

3. Der faktisch anonymisierte Mikrozensus – Das Scientific Use File –

Bei den vorliegenden Daten des Mikrozensus Scientific Use File 2001 handelt es sich um eine faktisch anonymisierte 70-Prozent-Substichprobe des Mikrozensus 2001. Für die faktisch anonymisierte Substichprobe gilt, dass sie als systematische Zufallsauswahl aus dem Originalmaterial des Statistischen Bundesamtes gezogen wurde. Das Material wurde dazu nach Bundesland, Regierungsbezirk, Gemeindegrößenklasse, Anzahl der Personen in Privathaushalten, Auswahlbezirksnummer und Nummer des Haushalts im Auswahlbezirk sortiert. Durch diese Anordnung ist gewährleistet, dass die Substichprobe hinsichtlich dieser Merkmale nur geringe zufallsbedingte Abweichungen aufweist. In diesem Prozess wurden Gemeinschaftsunterkünfte wie Wohnungen behandelt, das Merkmal „Zahl der Personen in Privathaushalten“ wurde nicht zur Ziehung der Substichprobe herangezogen. Bei Gemeinschaftsunterkünften erhielt jede Person eine eigene fortlaufende ("Haushalts"-)Nummer. Zur Stichprobenziehung diente die letzte (Einer-)Stelle der Haushaltsnummer. Alle Haushalte, deren Einerstelle einer von drei zufällig gezogenen einstelligen Zahlen entsprach, gelangten nicht in die Stichprobe (30% der Haushalte)³.

Weiterhin unterscheidet sich das Scientific Use File vom Originalfile des Mikrozensus dadurch, dass bestimmte Variablen im Scientific Use File, bedingt durch die Anonymisierung, in klassierter und vergrößerter Form verfügbar sind.

So sind im Scientific Use File die Regionalangaben Bundesland und eine vergrößerte Gemeindegrößenklasse, die nicht für die kleinen Bundesländer Bremen und Saarland ausgewiesen ist, enthalten. Mit Hilfe eines speziellen Codes der Variablen Gemeindegrößenklasse kann zwischen West- und Ost-Berlin (EF708=9) unterschieden werden. Auch bei den weiteren Variablen des Scientific Use Files werden je nach Bedarf Merkmalsvergrößerungen vorgenommen, so dass jede Ausprägung in den univariaten Verteilungen mindestens 5.000 Fälle in der Grundgesamtheit aufweist. Die Ausprägungen der Variable Nationalität werden so aggregiert, dass jede Nationalitätengruppe in der Grundgesamtheit wenigstens 50.000 Einwohner umfasst.

³ Das beschriebene Auswahlverfahren gilt grundsätzlich für alle Mikrozensus Scientific Use Files. Eine Ausnahme bilden die anonymisierten Files der Mikrozensen mit Fragen zur Wohnsituation, bei denen anstelle der Haushaltsnummern die Wohnungsnummern (die Endstellen der Wohnungsnummern) zur Stichprobenziehung herangezogen werden.

4. Aufbereitung des Rohdatenfiles zum Mikrozensus Scientific Use File 2001 und Überprüfung der Plausibilität des erstellten Systemfiles

4.1 Einlesen des Rohdatenmaterials und Erstellung des Systemfiles

Die vom German Microdata Lab bereitgestellten Setups für das Mikrozensus Scientific Use File 2001⁴ dienen zum Einlesen des Rohdatenmaterials und zum Erstellen von Systemfiles für die Programmpakete SAS, SPSS und STATA. Sie beinhalten Programmanweisungen zum Umwandeln von nicht-numerischen Variablen in numerische Variablen⁵, zur Definition von fehlenden Werten sowie zum Versehen der Variablen und ihrer Ausprägungen mit entsprechenden Labeln.

Bis zum Mikrozensus 1996 wurden bei der Aufbereitung der Rohdaten die nicht fortlaufenden Kategoriennummern (z.B. 1, 3, 4) in fortlaufende Nummern (z.B. 1, 2, 3) umgewandelt. Der Schlüssel 0 'ohne Angabe' wurde i.d.R. in 8 'ohne Angabe' umkodiert. Um die größtmögliche Nähe des Scientific Use Files zum Originalfile zu gewährleisten, wurden diese Umkodierungen im Mikrozensus 2001 wie auch in den vorhergehenden Mikrozensen 2000, 1999, 1998 und 1997 nicht vorgenommen. Die fehlenden Werte wurden möglichst in 0 umgewandelt. War dieser Wert schon vergeben, dann wurden die fehlenden Werte in 9 oder 8 (bzw. 99, 999, 98) umkodiert.

Um eine reibungsfreie Aufbereitung der Daten zu gewährleisten, empfiehlt es sich, die in den Setups für das jeweilige Programmpaket vorgesehenen Voreinstellungen zu berücksichtigen und je nach Leistung des verarbeitenden Systems die Setups schrittweise durchführen zu lassen. Sind die Daten eingelesen und ist das entsprechende Systemfile erstellt, können die Verteilungen ausgewählter Variablen des Scientific Use Files 2001 (vgl. Anhang A) zur Überprüfung des fehlerfreien Ablaufs der Datenaufbereitung herangezogen werden.

Als Referenz zur Plausibilitätsprüfung der Scientific Use Files dienen die in den Fachserien des Statistischen Bundesamtes Veröffentlichten Ergebnisse des jeweiligen Mikrozensus. Für den Mikrozensus 2001 sind dies die Reihe 3 (Haushalte und Familien), die Reihe 4.1.1 (Stand und Entwicklung der Erwerbstätigkeit) und die Reihe 4.1.2 (Beruf, Ausbildung und

⁴ Die SAS-, SPSS- und STATA-Setups können von Webseiten des German Microdata Lab: http://www.gesis.org/Dauerbeobachtung/Mikrodaten/Daten/Abteilungsdaten/Mikrozensen/mz_daten_index.htm heruntergeladen werden.

⁵ Im Mikrozensus Scientific Use File 2001 ist lediglich die Variable AF643, die in EF643 umgewandelt wird, davon betroffen.

Arbeitsbedingungen der Erwerbstätigen) der Fachserie 1⁶. Bei dem Vergleich der Ergebnisse des Scientific Use Files mit den entsprechenden Veröffentlichungen des Statistischen Bundesamtes ist darauf zu achten, dass die in diesen Reihen veröffentlichten Zahlen gewichtete, auf die Gesamtbevölkerung hochgerechnete Ergebnisse des Mikrozensus sind. Weiterhin beziehen sich die Zahlen auf nach bevölkerungs- oder erwerbstatistischen Konzepten definierten (Teil-) Populationen, die in der jeweiligen Reihe der Fachserien des Statistischen Bundesamtes erläutert werden⁷.

Im Folgenden werden die Modalitäten der Gewichtung und Hochrechnung der Scientific Use Files sowie die für den Vergleich mit den Veröffentlichungen des Statistischen Bundesamtes relevanten Bevölkerungsbegriffe und Abgrenzungen erläutert. Die konkreten Schritte und Ergebnisse der Plausibilitätsprüfung der aufbereiteten Mikrozensus Scientific Use Files 2001 beim German Microdata Lab werden im Abschnitt 4.4 dargestellt.

4.2 Gewichtung und Hochrechnung zum Vergleich mit den Veröffentlichungen des Statistischen Bundesamtes

Im Scientific Use File liegen vier Gewichte zur Kompensation der systematischen Ausfälle und des zufallsbedingten Stichprobenfehlers sowie zur Anpassung der Daten an die Bevölkerungsfortschreibung vor⁸:

EF750 Personenbezogener Gewichtungsfaktor

EF751 Haushalts-/Familien-Gewichtungsfaktor in der regionalen Anpassungsschicht

EF755 Personenbezogener Gewichtungs- und Hochrechnungsfaktor für die disproportionale 0,45%-Unterstichprobe/Ergänzungsprogramm (Der Auswahlatz dieser Unterstichprobe liegt auf Regierungsbezirksebene bei 0,4%, 0,6%, 0,8% oder 1% und beträgt im Bundesdurchschnitt 0,45%).

EF756 Gewichtungs- und Hochrechnungsfaktor im Regierungsbezirk (Haushalts- und Familienfaktor für das Ergänzungsprogramm)

Sollen die Ergebnisse von relativen Häufigkeiten des Scientific Use Files mit Veröffentlichungen des Statistischen Bundesamtes zum Mikrozensus 2001 verglichen werden, muss das entsprechende Gewicht des Mikrozensus (EF750, EF751, EF755, EF756) verwendet werden. Um auf 1% der Bevölkerung, also auf dem vollen Mikrozensus hochzurechnen, wird das Ge-

⁶ Vgl. die jeweilige Reihe der Fachserie 1, Statistisches Bundesamt, Wiesbaden 2002, Metzler-Poeschel, Stuttgart.

⁷ Vgl. den „Textteil“ der jeweiligen Reihe der Fachserie 1, Statistisches Bundesamt, Wiesbaden 2002, Metzler-Poeschel, Stuttgart.

⁸ Hierbei handelt es sich um die Gewichtungs- und Hochrechnungsfaktoren des Mikrozensus Originalfiles. Für die Scientific Use Files stehen keine spezifischen Gewichtungsfaktoren zur Verfügung.

wicht mit dem Kehrwert der Ziehungswahrscheinlichkeit für das Scientific Use File (100/70) multipliziert (hier am Beispiel personenbezogener Auswertungen):

```
COMPUTE XXXX=EF750 * 100/70
WEIGHT BY XXXX
```

Wie erwähnt, beziehen sich die in den Fachserien des Statistischen Bundesamtes veröffentlichten absoluten Zahlen auf die gewichteten, auf die Gesamtbevölkerung hochgerechneten Daten des Mikrozensus. Zur Herstellung der Vergleichbarkeit zwischen den absoluten Häufigkeiten des Scientific Use Files mit den Veröffentlichungen des Statistischen Bundesamtes müssen indes die Ergebnisse eigener Berechnungen noch zusätzlich mit 100 (Kehrwert der Ziehungswahrscheinlichkeit für den vollen Mikrozensus) multipliziert werden. Dies entspricht der Multiplikation des jeweiligen Gewichtungsfaktors mit 100/0,7.

4.3 *Bevölkerungsbegriffe und Abgrenzungen im Mikrozensus*

Zum Vergleich der Ergebnisse des Scientific Use Files mit den vom Statistischen Bundesamt veröffentlichten Zahlen sind die jeweiligen Bevölkerungs-, und Erwerbskonzepte zu berücksichtigen. Um solche Vergleiche zu erleichtern, gibt die folgende Tabelle einen Überblick über die Bevölkerungsbegriffe, deren Abgrenzung und die entsprechenden Gewichtungs- und Hochrechnungsfaktoren.

Oft ist eine Kombination der Bevölkerung am Ort der Hauptwohnung und der Bevölkerung in Privathaushalten sinnvoll, um die Bevölkerung am Ort der Nebenwohnung und in Anstalts- oder Gemeinschaftsunterkünften ausschließen zu können. In diesem Fall sind die Zahlen mit dem Haushalts- und Familienfaktor (EF751) hochzurechnen. In Veröffentlichungen des Mikrozensus wird diese Kombination jedoch nicht verwendet.

Tabelle 1: Bevölkerungsbegriffe, Abgrenzungen und Gewichtungen im Mikrozensus Scientific Use File 2001

Bevölkerungsbegriff Subpopulation	Filter zur Abgrenzung der Subpopulation	Gewichtungsfaktor
Wohnberechtigte Bevölkerung	-	EF750 Personen- Gewichtungsfaktor
Bevölkerung am Ort der Hauptwohnung	EF505=1 or EF505=2	EF750 Personen- Gewichtungsfaktor
Erwerbstätige	(EF505=1 or EF505=2) and EF504=1	EF750 Personen- Gewichtungsfaktor

Bevölkerungsbegriff Subpopulation	Filter zur Abgrenzung der Subpopulation	Gewichtungsfaktor
Erwerbstätige (EU-Arbeitskräfteerhebung, 0,45% Substichprobe)	(EF505=1 or EF505=2) and EF504=1 and EF738=1	EF755 Gewichtungsfaktor f. disprop. 0,45% Unter- stichprobe
Privathaushalte	EF506=1 and EF507=1	EF751 Haushalts-/Familien- Gewichtungsfaktor
Bevölkerung in Privathaushal- ten	EF506=1	EF751 Haushalts-/Familien- Gewichtungsfaktor
Bevölkerung am Familien- wohnsitz	EF508=1	EF75 Haushalts-/Familien- Gewichtungsfaktor
Familien	EF509=1 and EF508=1	EF751 Haushalts-/Familien- Gewichtungsfaktor
Nichteheliche Lebensgemein- schaften (verschieden geschlechtlich)	EF506=1 and EF515=1 and EF630=1-4	EF75 Haushalts-/Familien- Gewichtungsfaktor

4.4 Plausibilitätsprüfung des aufbereiteten Mikrozensus Scientific Use Files 2001

Die Plausibilitätsprüfung des Mikrozensus Scientific Use Files 2001 beim German Microdata Lab erfolgte durch den Vergleich des aufbereiteten Datensatzes mit den vom Statistischen Bundesamt veröffentlichten Tabellen in den entsprechenden Reihen der Fachserie 1. Dabei wurde nach den vom Statistischen Bundesamt verwendeten Bevölkerungs- und Erwerbskonzepten gewichtet und selektiert. Tabelle 2 enthält die Liste der zur Plausibilitätsprüfung herangezogenen Variablen, Angaben zu notwendigen Gewichtungen und Abgrenzungen sowie die entsprechenden Seitenzahlen der jeweiligen Datenreihe des Statistischen Bundesamtes.

Tabelle 2: Gewichtung, Abgrenzung und Vergleich mit den Fachserien des Statistischen Bundesamtes für ausgewählte Variablen des Mikrozensus Scientific Use File 2001

Fachserie 1 Reihe 3 (Haushalte und Familien)		
Gewichtung: Haushaltsgewicht		
Bevölkerung in Privathaushalten: EF506=1		
EF1:	Land der BRD	(Tab. 2.28, S.120)
EF30:	Alter (klassiert)	(Tab. 1.1, S.39)
EF32:	Geschlecht	(Tab. 1.1, S.39)
EF35:	Familienstand	(Tab. 1.1; S.39)
EF44:	Staatsangehörigkeit	(Tab. 2.19, S.102)
EF521:	Zahl der Personen in Privathaushalten	(Tab. 1.1, S.39)
EF541:	Haushaltstyp	(Tab. 2.24, S.112ff)
EF542:	Familientyp in den Haushaltstypen	(Tab. 2.24, S.112)
EF708:	Gemeindegrößenklasse	(Tab. 1.1, S.39)

Fachserie 1 Reihe 4.1.1 „Stand und Entwicklung der Erwerbstätigkeit“		
Gewicht: Personengewicht		
Bevölkerung am Ort der Hauptwohnung: EF505=1 oder EF505=2		
Erwerbstätige: EF505=1 oder EF505=2 und EF504=1		
Erwerbslose: EF505=1 oder EF505=2 und EF504=2 oder EF504=3		
EF127:	Stellung im Beruf	(Tab. 1.2; S.43)
EF129:	Wirtschaftszweige	(Tab. 3.4; S.65)
EF138:	Vollzeit/Teilzeittätigkeit	(Tab. 3.17; S.102)
EF139:	Grund für Teilzeittätigkeit	(Tab. 3.17; S.102)
EF213:	Grund der Arbeitsuche bei NEWT	(Tab. 4.5, S. 112)
EF216:	Tätigkeit als Selbständiger/Arbeitnehmer	(Tab. 4.4, S.110)
EF217:	Art der gesuchten Tätigkeit	(Tab. 4.4, S.110)
EF239:	Dauer der Arbeitsuche	(Tab. 4.3, S.108)
EF248:	1. Aktive Methode der Arbeitsuche (Arbeitnehmer)	(Tab. 4.4, S.110)
EF287:	Allgemeinbildender Abschluss	(Tab. 3.16.1, S.94)
EF289:	Beruflicher Ausbildungsabschluss	(Tab. 3.16.1, S.94)
Fachserie 1 Reihe 4.1.2 (Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen)		
Gewichtung: a) Personengewicht (EF750); b) Personengewichts für die 0.45%-Unterstichprobe (EF755)		
Bevölkerung am Ort der Hauptwohnung und Erwerbstätige: EF505=1 oder EF505=2 und EF504=1		
a):		
EF128:	Beruf	(Tab. 11, S.77)
b):		
EF147:	Samstagsarbeit	(Tab. 14, S. 100)
EF154:	Spätschicht	(Tab. 16, S.110)

Die Programmanweisungen für die durchgeführten Plausibilitätsprüfungen sind als SPSS-Syntax im Anhang B aufgeführt.

Bedingt durch die Ziehung der 70%-Stichprobe ergeben sich Abweichungen zwischen den Häufigkeiten der Variablen des Scientific Use Files und den in Fachserien des Statistischen Bundesamtes veröffentlichten Zahlen. Die Mehrzahl der oben aufgeführten Variablen weichen i.d.R. nur in geringem Maße (meist 0% bis 1%, max. 5%) von den veröffentlichten Daten ab. Abweichungen über 5% wurden bei Variablen festgestellt, deren Merkmalsausprägungen mit sehr kleinen Fallzahlen besetzt sind (z.B. Staatsangehörigkeit und Wirtschaftsbereiche).

Anhang A: Randverteilung ausgewählter Variablen des Scientific Use Files 2001

Bevölkerungskonzept	Bevölkerung in Privathaushalten	Bevölkerung am Ort der Hauptwohnung	Bevölkerung am Familienwohnsitz	Erwerbstätige Bevölkerung am Ort der Hauptwohnung	Fallzahlen
	<i>Haushaltsgewicht⁹ select if (ef506=1)</i>	<i>Personengewicht¹⁰ select if (ef505=1 or ef505=2)</i>	<i>Haushaltsgewicht select if (ef508=1)</i>	<i>Personengewicht select if (ef505=1 or ef505=2) and (ef504=1)</i>	<i>ohne Selektion und ohne Gewichtung</i>
Insgesamt	826029	823029	815766	368648	503961
EF32: Geschlecht					
Männlich	401973	402044	396442	206427	243133
Weiblich	424056	420985	419325	162222	260828
EF35 Familienstand					
Ledig	317777	314123	312198	115336	192972
Verheiratet	404753	402074	400825	222253	245297
Verwitwet	61212	64446	60809	5902	39854
Geschieden	42286	42386	41934	25156	25838
EF30: Alter (Gruppiert)					
unter 20	172329	170357	170739	13430	103296
20-25	47679	45354	45309	30252	28266
25-30	46701	45418	45215	34156	27572
30-35	62703	61890	61535	49531	37573
35-40	69690	69027	68752	56466	42028
40-45	64015	63445	63166	52201	38748
45-50	58790	58088	57928	46936	35561
50-55	53909	53325	53098	39660	32532
55-60	50896	50237	50004	29000	30799
60-65	61096	60386	59985	12885	37302
65-70	45269	44902	44530	2542	27734
70-75	46028	45651	45049	1172	28336
76 und älter	55336	54949	50455	418	34214

⁹ Gewichtet mit Haushaltsgewicht (compute hhgew=ef751*100/70; weight by hhgew)¹⁰ Gewichtet mit Personengewicht (compute persgew=ef750*100/70; weight by persgew)

Bevölkerungskonzept	Bevölkerung in Privathaushalten	Bevölkerung am Ort der Hauptwohnung	Bevölkerung am Familienwohnsitz	Erwerbstätige Bevölkerung am Ort der Hauptwohnung	Fallzahlen
EF44: Staatsangehörigkeit					
Bosnien-Herzegowina	1816	1846	1813	862	858
Frankreich	1277	1309	1249	699	653
Griechenland	3831	3868	3812	1981	1793
Grossbritannien	1268	1338	1240	776	627
GUS	3553	3650	3527	1368	1877
Italien	6969	7102	6898	3762	3389
Kroatien	2952	2980	2924	1753	1409
Niederlande	1177	1257	1152	635	556
Österreich	1872	1953	1811	1130	930
Polen	2644	2774	2619	1429	1434
Portugal	1434	1461	1430	901	652
Rumänien	996	1040	990	626	553
Übrige EU ¹¹	1067	1095	1034	561	527
Serbien und Monteneg.	4416	4577	4391	1879	2147
Slowakei etc. ¹²	833	868	813	513	435
Spanien	1239	1264	1212	673	606
Türkei	23727	24038	23626	8244	10971
Übriges Europa ¹³	3073	3139	3018	1269	1540
Marokko	801	830	798	249	407
Sonstiges Afrika	1427	1519	1410	624	722
USA	1260	1302	1226	649	661
Sonstiges Amerika	822	882	810	365	437
Iran	962	963	942	422	472
Sonst. Naher Osten ¹⁴	1764	1878	1757	457	833
Vietnam	754	781	743	311	354
Sonst. Südasien ¹⁵	2159	2263	2136	733	1059
Ostasien ¹⁶	1184	1221	1162	469	569
Übrige Welt, Staatenlos	843	892	838	354	424
Deutscher ¹⁷	749907	744937	740386	334954	467066
EF338: Überw. Lebensunterhalt					
Erwerbs-/Berufst.	341406	337097	336099	337097	206460
Arbeitslosengeld/-Hilfe	28502	28414	28348	1477	16869
Rente, Pension	181776	185033	180263	5238	114430
Unterhalt	247950	244019	245390	21158	149130
Eigenes Vermögen	3811	3725	3653	415	2373
Sozialhilfe	13181	15047	13123	1102	8492
Leist. Pflegevers.	552	1163	549	51	722
Sonst. Unterstützung	8851	8531	8343	2111	5485

¹¹ Belgien, Dänemark, Finnland, Irland, Luxemburg, Schweden¹² Tschechische Republik, Ungarn¹³ Schweiz, Sonstiges Ost- und Mitteleuropa, Sonstiges Westeuropa¹⁴ z.B. Libanon, Jordanien, Syrien, Irak, Israel¹⁵ z.B. Afghanistan, Indien, Kambodscha, Laos, Pakistan, Thailand, Sri Lanka¹⁶ z.B. China, Hongkong, Indonesien, Japan, Korea, Macao, Philippinen¹⁷ ohne weitere Staatsbürgerschaft

Bevölkerungskonzept	Bevölkerung in Privathaushalten	Bevölkerung am Ort der Hauptwohnung	Bevölkerung am Familienwohnsitz	Erwerbstätige Bevölkerung am Ort der Hauptwohnung	Fallzahlen
EF504: Erwerbstyp					
Erwerbstätig	373252	368648	367378	368648	226104
Erwerbslose (EU) ¹⁸	31570	31489	31388	0	18611
Sonst. Erwerbslose	5898	5867	5842	0	3488
Nichterwerbspersonen	415309	417025	411158	0	255758
EF127: Stellung im Beruf					
Selbst. ohne Beschäftigten.	18533	18142	18111	18142	11272
Selbst. mit Beschäftigten	18281	17984	17894	17984	11134
Mithelf. Familienangehöriger	4522	4453	4463	4453	2771
Beamter, Richter	19844	19288	19325	19288	12267
Angestellter	179839	176096	176487	176096	109784
Heim-(Arbeiter)	112067	112418	111359	112418	66856
Kaufm.-techn. Auszubildender	8109	7840	7901	7840	4937
Gewerbl. Auszubildender	7936	7831	7802	7831	4829
Zeit-/Berufssoldat	2077	2346	2030	2346	1092
Grundwehrdienstleist.	979	1225	979	1225	504
Zivildienstleistende	1066	1026	1027	1026	658
Entfällt	452777	455737	448388	0	277857
EF372: Höhe d. Nettoeinkommens im April					
Unter 153 Euro	23195	23898	23014	4756	14571
153 b.u.307 Euro	36632	36472	36191	15543	22168
307 b.u.511 Euro	58103	57123	56853	28856	35285
511 b.u.716 Euro	67205	66177	66090	25634	40685
716 b.u.920 Euro	63553	63083	62883	27638	38557
920 b.u. 1.125 Euro	72387	72156	71707	38391	44058
1.125 b.u.1.278 Euro	60685	60581	60105	37993	36967
1.278 b.u.1.534 Euro	66811	66568	65997	46569	40691
1.534 b.u.1.790 Euro	48772	48647	48055	37281	29831
1.790 b.u.2.045 Euro	30368	30565	29914	24319	18794
2.045 b.u.2.301 Euro	20445	20630	20023	16584	12826
2.301 b.u.2.556 Euro	14381	14632	14046	11736	9149
2.556 b.u.2.812 Euro	9846	9860	9582	7952	6211
2.812 b.u.3.068 Euro	6918	6927	6737	5707	4349
3.068 b.u.3.323 Euro	5485	5418	5305	4594	3426
3.323 b.u.3.579 Euro	3579	3510	3476	3071	2212
3.579 b.u.3.835 Euro	2626	2568	2536	2232	1645
3.835 b.u.4.090 Euro	2613	2575	2537	2212	1619
4.090 b.u.5.113 Euro	3934	3838	3782	3408	2442
5.113 b.u. 6.136 Euro	2160	2050	2036	1873	1325
6.136 b.u. 7.669 Euro	1201	1125	1116	1028	736
7.669 b.u. 10.226 Euro	893	848	844	771	547
10.226 b.u. 17.895 Euro	607	582	579	524	374
17.895 und mehr	606	563	564	511	373
Selbst. Landwirt	2674	2665	2657	2663	1664
Kein Einkommen	189389	187423	188571	1196	113536
Keine Angabe	30961	32545	30565	15605	19920

¹⁸ sofort verfügbar (EU-Definition)

Bevölkerungskonzept	Bevölkerung in Privathaushalten	Bevölkerung am Ort der Hauptwohnung	Bevölkerung am Familienwohnsitz	Erwerbstätige Bevölkerung am Ort der Hauptwohnung	Fallzahlen
EF287: Höchster allg. Schulabschluss					
Haupt-/Volksschule	306060	307085	304508	131172	187020
POS	49910	49590	49644	37273	30085
Realschulabschluss	128109	126673	126392	87566	78417
Fachhochschulreife	25837	25273	25209	18882	15713
Abitur/Fachabitur	109579	104854	104709	73467	66338
Angabe fehlt	5458	5469	5336	2952	3337
Entfällt	201076	204084	199967	17338	123051
EF289: Höchster berufl. Ausbildungsabschluss					
Anlernausbildung	11387	11468	11315	4925	6983
Berufsvorbereitungsjahr	1525	1525	1514	766	933
Lehrausbildung	303482	302249	301002	179131	185670
Berufsfachschule	20099	19862	19779	12593	12354
Meister, Techniker	42921	42475	42306	29915	26334
Fachschule ehem. DDR	10450	10352	10386	6535	6300
Fachhochschule	24566	23964	23892	17993	14992
Hochschulabschluss	39479	38200	38051	29533	23948
Promotion	5997	5671	5640	4361	3651
Angabe fehlt	11404	11468	11117	7010	6977
Entfällt	354719	302249	350763	75886	215819
EF138: Vollzeit-/Teilzeittätigkeit					
Vollzeittätigkeit	296846	293736	292129	293736	179532
Teilzeittätigkeit	76406	74913	75249	74913	46572
Entfällt	452777	454380	448388	0	277857

Anhang B: SPSS-Syntax für die Plausibilitätsprüfungen des Scientific Use Files 2001

*** Plausibilitätsprüfung: Vergleich mit den Fachserien des Statistischen Bundesamtes MZ2001.**

get file = ['SPSS-Datenfile'].

*Bildung des Personengewichts.

compute persgew=ef750*(100/70).

*Bildung des Haushalts- und Familiengewichts.

compute hhgew=ef751*(100/70).

title 'Plausibilitätsprüfung'

subtitle 'Fachserie 1, Reihe 3 (Haushalte und Familien)'.

Weight by hhgew.

recode ef30

(0 thru 24=1)(25 thru 34=2)(35 thru 44=3)(45 thru 54=4)

(55 thru 64=5) (65 thru 74=6) (75 thru 95=7).

*Filter: Bevölkerung in Privathaushalten.

select if (ef506=1).

fre ef1 ef30 ef32 ef35 ef521 ef541 ef542 ef708.

*Filter: Bevölkerung in PHH und Haushaltsbezugsperson.

temporary.

select if (ef506=1) and (ef507=1).

fre ef44.

Weight off.

*****.

subtitle 'Fachserie 1, Reihe 4.1.1 (Stand und Entwicklung der EWT)'.

weight by persgew.

*Filter: Bevölkerung am Ort der Hauptwohnung.

temp.

select if (ef505=1 or ef505=2).

fre ef127.

*Filter: Bevölkerung am Ort der Hauptwohnung, nur Erwerbslose.

temp.

select if (ef505=1 or ef505=2) and (ef504=2 or ef504=3).

fre ef216 ef217 ef239 ef248.

*Filter: Bevölkerung am Ort der Hauptwohnung, nur Erwerbslose

(arbeitsuchend).

temporary.

select if (ef505=1 or ef505=2)and(ef504=2 or ef504=3)and(ef208=1).

fre ef213.

recode ef129 (11 thru 50=1)(101 thru 145=2)(151 thru 372=3)

(401 thru 454=4)(501 thru 555=5)(601 thru 642=6)

(651 thru 672=7)(701 thru 748=8)(751 thru 753=9)

(801 thru 990=10).

```

value label ef129 1 'Land- und Forstwirtschaft, Fischerei'
                2 'Bergbau'
                3 'Verarbeitendes Gewerbe'
                4 'Energie- und Wasserversorgung'
                5 'Handel, Gastgewerbe'
                6 'Verkehr, Nachrichten'
                7 'Kredit- und Versicherungsgewerbe'
                8 'Grundstueck etc'
                9 'Oeffentliche Verwaltung'
               10 'Oeffentliche und private Dienstleistungen'.

*Filter: Bevoelkerung am Ort der Hauptwohnung, nur Erwerbstaetige.
temporary.
select if (ef505=1 or ef505=2)and(ef504=1).
fre ef129.

*Filter: Bevoelkerung am Ort der Hauptwohnung,nur Abhaengige Erwerbstaetige
mit Auszubildenden.
temporary.
select if
(ef505=1 or ef505=2)and(ef504=1)and(ef127>=4 and ef127<=8)).
fre ef138 ef139 ef287 ef289.

Weight off.

*****.

subtitle 'Fachserie 1, Reihe 4.1.2 (Beruf, Ausbildung und Arbeitsbedingun-
gen der Erwerbstaetigen)'.

weight by persgew.

*Filter: Bevoelkerung am Ort der Hauptwohnung und Erwerbstaetige.
temporary.
select if ((ef505=1 or ef505=2)and(ef504=1)).
fre ef128.

weight off.

* Bildung des Personengewichts für die 0.45%-Unterstichprobe.
comp pergew_U=ef755*(100/70).
weight by pergew_u.

*Filter: Bevoelkerung am Ort der Hauptwohnung und Erwerbstaetige.
temporary.
select if ((ef505=1 or ef505=2)and(ef504=1)).
fre ef147 ef154.
Weight off.

```

Anhang C: Weiterführende Literaturhinweise zum Mikrozensus

Informationen zum Mikrozensus und seiner Konzeption:

- Emmerling, D.; Riede, T. (1997): 40 Jahre Mikrozensus. *Wirtschaft und Statistik* (3): 160-174.
- Hartmann, P. (1989): Der Mikrozensus als Datenquelle für die Sozialwissenschaften. *ZUMA-Nachrichten* 24: 6-25.
- Heidenreich, H.-J. (1994): Hochrechnung des Mikrozensus ab 1990. In: Gabler, S./Hoffmeyer-Zlotnik, J./Krebs, D. (Hrsg.): *Gewichtung in der Umfragepraxis* (S. 112-123). Opladen: Westdeutscher Verlag,
- Lüttinger, P.; Riede, T. (1997): Der Mikrozensus: amtliche Daten für die Sozialforschung. *ZUMA-Nachrichten* 41: 19-43.
- Meyer, K. (1994): Zum Auswahlplan des Mikrozensus ab 1990. In: Gabler, S./Hoffmeyer-Zlotnik, J./Krebs, D. (Hrsg.): *Gewichtung in der Umfragepraxis* (S. 106-111). Opladen: Westdeutscher Verlag.
- Statistisches Bundesamt (1997): *Arbeitsunterlagen zum Mikrozensus. Das Erhebungsprogramm des Mikrozensus seit 1957*. Wiesbaden .
- Statistisches Bundesamt (1989): *Inhaltliche Fragen bevölkerungstatistischer Stichproben am Beispiel des Mikrozensus. Bericht zur Konferenz vom 21. und 22. Oktober 1988. Heft 10 der Schriftenreihe Ausgewählte Arbeitsunterlagen zur Bundesstatistik*. Wiesbaden.
- Schmidt, S. (2000): *Erwerbstätigkeit im Mikrozensus. Konzept, Definition, Umsetzung*. ZUMA-Arbeitsbericht 2000/01.

Zur Weitergabe faktisch anonymisierter Einzeldaten:

- Müller, W.; Blien, U.; Knoche, P.; Wirth, H. (1991): *Die Faktische Anonymität von Mikrodaten*, (Band 19 der Schriftenreihe Forum der Bundesstatistik). Stuttgart: Metzler-Poeschel.
- Wirth, H. (1992): Die faktische Anonymität von Mikrodaten: Ergebnisse und Konsequenzen eines Forschungsprojektes. *ZUMA-Nachrichten* 30: 7-42

Zur Fehlerrechnung im Mikrozensus:

Krug, W.; Nourney, M.; Schmidt, J.: (1994): Wirtschafts- und Sozialstatistik. Gewinnung von Daten., München: Oldenbourg (3. völlig Neubearb. Auflage) [s. insbes. Seite 242-246].

Statistisches Bundesamt (2001): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.1, Stand und Entwicklung der Erwerbstätigkeit 2001 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Ergebnisse des Mikrozensus 2001:

Statistisches Bundesamt (2002): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 3, Haushalte und Familien 2001 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Statistisches Bundesamt (2002): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.1, Stand und Entwicklung der Erwerbstätigkeit 2001 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Statistisches Bundesamt (2002): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.2, Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen 2001 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Vergleichbarkeit der Mikrozensen:

Kollinger, N.; Kohlmann, A. (2003): Vergleich der Mikrozensen 1998 und 2000. ZUMA-Methodenbericht 2003/08.

Kollinger, N.; Shahla, H. (2004): Vergleich der Mikrozensen 2000 und 2001. ZUMA-Methodenbericht 2004/09.

Lehnert, D.; Kohlmann, A. (2003): Vergleich der Mikrozensen 1997 und 1998. ZUMA-Technischer Bericht 2003/02.

Schimpl-Neimanns, B.; Schmidt, K. (1998): Vergleichbarkeit der faktisch anonymisierten Mikrozensen 1989, 1991, 1993, 1995. ZUMA-Technischer Bericht 1998/09.

Staudenmaier B.; Granato, N.; Schmidt, S. (2000): Vergleich der Mikrozensen 1995 und 1996. ZUMA-Technischer Bericht 2000/08.

Staudenmaier B./Schmidt, S. (2000): Vergleich der Mikrozensen 1996 und 1997. ZUMA-Technischer Bericht 2000/09.

Publikationen, die auf den Scientific Use Files der Mikrozensen basieren, finden sich auf der Webseite http://www.gesis.org/Dauerbeobachtung/Mikrodaten/Publikationen/MZ_Literatur.htm.

Darüber hinaus sind Veröffentlichungen in der Zeitschrift 'Wirtschaft und Statistik' des Statistischen Bundesamtes zu finden. Ein umfassendes Informationsangebot bietet das German Microdata Lab, ZUMA an: <http://www.gesis.org/Dauerbeobachtung/Mikrodaten/index.htm>.